BIO Web of Conferences **4**, 00014 (2015)
DOI: 10.1051/bioconf/20150400014

From *neonts* to *filionts* and their progenies...

Biodiversity draws its origins from the origins of life itself

Marie-Christine Maurel[a]

Institut de Systématique, Évolution, Biodiversité (ISYEB - UMR 7205 – CNRS, MNHN, UPMC, EPHE), Muséum national d'Histoire naturelle, Sorbonne Universités, 45 rue Buffon, CP. 50, 75005 Paris, France

"Should the question of the priority of the egg over the chicken, or of the chicken over the egg disturb you, it is because you assume that animals were originally what they are to-day; what madness!"

The dream of d'Alembert, 1760, Denis Diderot

Abstract. Since the origins, different modes of life exist. Indeed it is difficult to imagine how things could have been constructed by a single event that would have marked the birth of life. A brief historical background of the subject will be presented here. Then, we shall wander through the history of the discoveries that have led the way of experimental science over the last 150 years and shall dwell on one of the major paradigms, that of the RNA World. This paradigm illustrates the path followed by the living that tries all possible means when expressed at the molecular level as simple free molecule leading to RNA fragments and to modern subviral particles such as viroids then to ramifications and compartments. Between the ancient and the modern RNA world we underline how environment talks to molecules leading to the global living tissus which establish the ground of biodiversity.

1. Introduction

It is common knowledge that a population of living beings is composed of individuals that are never formed *de novo* starting from surrounding compounds but are the result of heredity with variations. Hence, how can one explain the constitution of living beings starting from a "soup" of primordial molecules?

This is the question raised by studies on the origins of life.

"I conjecture that the origin of life and the origin of problems coincide". This statement by Karl Popper [1] definitely placed, as of pre-Darwinian evolution, the role of natural selection as adaptive factor based on competition for the recovery of biologically interesting molecules.

The biotic capacity of the Earth made it possible for such a particular environment and linked to it, to develop competing molecular variants to exploit these resources. Competition in chemistry and

[a] Corresponding author: marie-christine.maurel@upmc.fr

biochemistry is most frequently defined in terms of rate of the reaction and of the dynamics that allow evolution of the reactions. Structural configurations and plasticity are required, and one observes that as a result of physical perturbations, cyclic behaviours and/or a certain resilience are produced that are favourable for the elaboration of perennial molecules. This is probably how the biochemical identity of variants was constituted, and acts of reorganisation absorbed the perturbations of the environment as transmissible informations.

Hence the question: starting from an initial chaos of molecules without apparent clear links between them, how were the first molecules formed transporting genetic information and how were the first catalysts formed?

1.1 Competition and community of origin define the bases of a Darwinian evolution as of the origins

The complicated system of heredity based on identical duplication of chromosomes composed of DNA or RNA cannot have been the first. It is probably better to consider that at the origins of life, the environmental conditions corresponded to what was permanent, that is the macroscopically defined conditions of the primitive soup. The random state must have resided on the side of the condensation of molecules, of assembly into coacervates (from the latin coacerver: to collect, to cluster) as they were defined by Oparin [2, 3]. We must mention that molecular clusters might have existed without globular coacervation, but just as templates. At the level of coacervates the first globules must have differed from one another in their composition (composed of amino acids, that is proteneoid microspheres, of arabic gum, or lipids as in liposomes) and in their contents (catalysts, informative molecules, signals, . . .); the daugther proto-cells deriving from the division of the same coacervate did not share all the components of the mother proto-cell with a precision similar to the one that will appear later in the reproduction of chromosomes. Errington recently observed (2014) and established the morphological comparison of the fusion/fission process of the *L-Bacillus subtilis* (bacteria devoid of cell wall) with coacervates [4]. Both produce offsprings in the form of “buds”, kinds of blebs perfectly observable by electron microscopy.

It is difficult to imagine how things could have been constructed by a single operation that would have marked the birth of life. After the assembly of powerful molecules, there no doubt existed a period during which populations of more or less stable globules enrolling efficient molecules undergoing random and unequal divisions producing neo-formed individuals and designated more or less ephemerally as *neonts*, preceded similar individuals yet capable of reproduction by division, the *filionts* [5, 6]. It is likely that the conditions required for the formation of populations of neonts arose on several occasions followed by extinctions or primers of evolution but without continuation.

A long process of trials and errors, a kind of natural proto-selection was put into place before the elements capable of transmitting a specific organization became stabilized. During this process, the capacity of adaptive reactions became more acute and specialized with the establishment of exchanges with the environment and the elaboration of reactions when faced with these changes.

Consequently, acquiring a specificity transmissible from one generation to the next must have constituted an important step in the origins of life. Once acquired, maintaining it in spite of fluctuations in the environment must have constituted a requirement of individual adaptation. Thus it is very likely that as of the origins of life, the double question of the transmission of a specific organisation and of a mode of individual adaptation that is not transmissible must have been resolved by the living beings.

The solution became stabilized in the form of specialisation and complementarity between the constituents of the cell capable of complete duplication, and the other components some of which are more specifically aimed at establishing exchanges with the external milieu. Once this specialization acquired, living organisms can be defined as all organism capable of evolving under the influence of natural selection.

The existence of mixed transitional populations of newly formed molecules and proto-cells, the *neonts*, and of slightly different populations but capable of reproduction by division, the *filionts* that appeared by reproduction of pre-existing ancestors lead one to wonder about the ecological conditions of their evolution which ultimately led to the formation of the "planetary living tissue" that is the loam of biodiversity.

More or less regular cycles of hydration-dryness might have facilitated the multiplication of evolutionary trials, introducing repetition by reiteration leading to short-lived filionts most certainly destined to rapid extinction. However, should reproduction be likely if the probability of reproduction is linked to a particular biochemical composition, a process of natural selection is put into place since a certain probability exists that at least one of the two filionts inherited the chemical personality that will have simplified the reproduction of its parent.

1.2 Biodiversity acquires its origins from the origins of life itself

The transition from the inert to the living thus shows that from the start, one can be alive in different ways. In 1998, Carl Woese insisted on the necessity of not representing the universal ancestor of all living organisms as a precise and well-defined strain. He proposed the hypothesis of multiple lines poorly individualized between which many lateral gene transfers would have been possible. Thus LUCA (the last universal common ancestor), a population of living beings, presumably of prokaryote-type, corresponds to the minimum total number of essential genes. In prokaryotes, there are 300 to 10 000 genes. This already represents considerable original diversity. Hence, LUCA must be considered as a community of primordial organisms themselves deriving from an extraordinary original molecular diversity that as of the origins performed biochemical acts more or less efficiently: energetic acts, catalytic acts, replicative acts, signaling and interactive acts enabling elementary feedback, conquering ecological niches, crucibles of mineral compartments, then of organic compartments. Selection supposes the production of diversity prior to the filter of experience. Natural selection functions only between what is viable, that is everything which for at least a certain time, interacts with the environment. Thus, the moment it begins, life is faced with survival. From the non-living to the living, gradually and based on several dimensions at various levels, life became organized as we know it through the fossil records and as observed today.

These multiple origins are compatible with the originalities of the living known today as biodiversity. The discoveries of Nature, as pointed out by Lucrece occur "... after wandering through the infinity of time, trying all types of unions, all possible movements, finally reaching these assemblies, that when suddenly united, are at the origin of these large objects, ..., *the living species*" [7].

2. A brief historical background

In the following section, we shall wander through the history of the discoveries that have led the way of experimental science over the last 150 years and shall dwell for some time on one of the major paradigms, that of the RNA World which illustrates the path followed by the living that tries all possible means when expressed at the molecular level as simple virus or subviral particles such as viroids.

Life first settled on the primitive Earth less than a billion years after the formation of the Solar System. How did chemical transformations allow the passage from mineral matter to simple organic matter capable of forming the elementary bricks of biomolecules, that are the amino acids of proteins, nitrogen bases, sugars, phosphate, the constituants of nucleotides, components of the nucleic acids DNA and RNA, fatty acids and lipids? How did these elementary building blocks organize themselves and put into place the first autoreplicative molecules related to our nucleic acids? What were the first catalysts involved in bringing together the first metabolic links, allowing them to function? What happened during

this first billion years of the Earth's history that led to the advent of organisms preserved in stromatolites similar to present-day bacteria?

Chemical evolution led the simplest elements, hydrogen, carbon, nitrogen, oxygen, sulfur, phosphorous, etc. to combine and form more complex organogenous compounds: methane, carbon monoxyde, carbon dioxyde, water vapor, etc. These molecules that were present in the atmosphere of the primitive Earth are still synthesized today in the interstellar space.

When a gaseous mixture of this type is rapidly heated, and then submitted to an electric discharge or to another form of energy (mechanics, light, radioactivity etc.) complex organic compounds form spontaneously. These have been experimentally investigated and constitute the basis of what is known as prebiotic chemistry.

It is interesting to note that laboratory chemical synthesis of biologically important organic compounds were obtained during the late 19th and early 20th centuries, but none of them were related to the study of the origins of life.

The most notable facts are the following. In 1828, Friedrich W′ohler synthesized urea starting from ammonium cyanide, thus overcoming what was at the time considered as an insurmountable barrier between organic and mineral matter [8]. In 1850, Adolph Strecker produced alanine by mixing hydrogen cyanide, acetaldehyde and ammonia, much before this amino acid was identified in proteins in 1874. A few years later, in 1861, Aleksandr Butlerov polymerized formaldehyde (by the so-called formose reaction) producing sugars including pentoses [9]. In 1913, Walter Lôb produced the amino acid glycine by submitting a mixture of carbon dioxide, ammonia and water vapor to an electric discharge [10], and finally in 1926 David Davidson and Oskar Baudish obtained uracil, a component of nucleic acids, from urea [11].

None of the authors of these remarkable studies was interested in the origins of life. For instance, Walter Lôb who obtained glycine, was trying to understand how atmospheric nitrogen was assimilated by plants. His article appeared nearly unnoticed and even today, it is largely unknown.

In parallel, as of 1859, the impact of Charles Darwin on our understanding of the origin of species was immense. A few years later (1871), the founder of the theory of evolution presented, in a letter to Joseph Hoocker, his speculation on the chemical origin of life:

> "It is often said that all the conditions for the first production of a living organism are now present, which could ever have been present. – But if (and oh! what a big if!) we could conceive in some warm little pond, with all sorts of ammonia and phosphoric salts, – light, heat, electricity, etc., present that a protein compound was chemically formed ready to undergo still more complex changes, at the present day such matter would be instantly devoured, or absorbed, which would not have been the case before living creatures were formed" [12].

In 1924, the crucial contribution of Alexander Ivanovitch Oparin to our understanding of the origins of life followed the same trend. The Soviet scientist stressed the gradual aspect of these processes, the importance of the environment and of the conditions in which these "building blocks of life" and proto-cells were formed. With this new step, Oparin opened the way to the experimental approach in this area that was often considered at the time as philosophical (in a partial and restrictive sense) or belonging exclusively to the domain of ideas. Oparin updated all the data and by placing the problem of the origins in an environmental, geophysico-chemical and paleontological setting, he initiated nearly 100 years of scientific and experimental research in this domain. Calling on the most contemporary biochemical data available, he layed out the problems to be solved to understand the sequence of extraordinary events that were to lead to the origins of life.

In 1929, the English biologist and geneticist John Burdon Sanderson Haldane proposed the same type of scenario as his Sovietic homologue, and this even before being aware of the latter's study that appeared only in Russian. The compounds dissolved in oceans would have constituted the "prebiotic soup" in which organic molecules possibly capable of self-assembly would have appeared. Proto-cells

or coacervates would then have concentrated from this primitive "broth" the molecules capable of developing a proto-metabolism, a true bridge between the chemical world and the living world. The construction of the first cells would thus have occurred in parallel with the birth of heterotrophic metabolism.

3. The first experimental stages

The hypothesis of Oparin and Haldane was experimentally tested in the laboratory of the Nobel prize winner (1934) Harold Clayton Urey in Chicago who proposed a model of primitive reducible atmosphere rich in methane. In 1953, Stanley Miller reproduced in Urey's laboratory this atmosphere and submitted a mixture of methane, ammonia, hydrogen and water to an electric discharge, simulating thunderbolts [13]. This experiment led to the formation of organic compounds including 5 amino acids in all aspects identical to those found in living organisms. Amino acids can therefore be produced "spontaneously" from organic matter without the intervention of enzymes.

A few years later, in 1961, Juan Orò obtained purines, mostly adenine, by polymerizing 4 molecules of hydrogen cyanide (HCN) under the effect of UV rays in the presence of ammonia [14, 15]. This was followed by a series of syntheses, mainly of adenine and nicotinamide, through photochemical rearrangements of hydrogen cyanide [16, 17].

We must recall here that in the 40's and based on studies of Daniel Berthelot and of Henry Gaudechon whose aim was to synthesize sugars *in vitro* starting from water and carbon dioxide submitted to the rays of a quartz lamp of mercury vapor [18], Alexandre Dauvillier and Etienne Desguin developed a photochemical hypothesis of original photosyntheses "performed on the surface of primitive sea waters" [19]. Innumerable experiments were listed by these authors, among which the synthesis of formamide, H_2NCOH, and we know today that it is an unavoidable agent of prebiotic synthesis of nitrogen-based nucleic acids, DNA and RNA. Moreover, formamide is found in space and in the environment of solar-type stars that are being formed.

Purine nitrogenous bases adenine (A), guanine (G) and pyrimidine nitrogenous bases cytosine (C) and uracil (U), have recently been produced by heating formamide in the presence of mineral catalysts and of UV photons. The four RNA bases could thus have been formed from formamide on a primitive Earth devoid of ozone layer and under the direct influence of UV rays.

The majority of present-day protein enzymes are aided by co-factors, most of which are ribonucleotidic co-enzymes. Examples are nicotinamide adenine dincleotide (NAD) that can undergo phosphorylation (NADP), riboflavines and FAD (flavine adenine dinucleotide) and FMN (flavine mononucleotide). Prebiotic synthesis of nicotinic acid derivatives were obtained from ethylene and ammonia by Friedmann et al. in 1971 [20]. Flavine is a tri-cyclic heterocycle, based on a pteridine nucleus. Purines, derivatives of nicotinamide and pteridines are obtained by photochemical rearrangement of HCN known to be abundantly present everywhere in the universe. Co-factors and co-enzymes are considered molecular fossils of an ancient RNA world, and a common origin starting from HCN can be retained. The structural similarity of purines and pteridines (where one pyperazine cycle replaces the imidazole moiety of the purine) is also very interesting (Fig. 1). Pteridine is found today in pigments, flavines and folates, that are indispensible biochemical co-factors known to be sensitive to light. Folates (precursors of tetrahydrofolate, THF) also participate in the biosynthesis of purine nucleotides. Without going into details of the biochemical paths [21], it is important to stress the metabolic relationship between all these compounds sensitive to light and that today participate to various extents in the biological oxydo-reduction reactions that produce energy.

On the other hand, a complex mixture of sugars is formed when formaldehyde is shaken in the presence of chalk or lime. The synthesis of these compounds first elaborated by Butlerow in 1861 occurs via a series of autocatalytic steps described in 1959 by Breslow [22]. However, this reaction presents a certain number of problems: if one attempts to obtain a sugar of biological interest, one

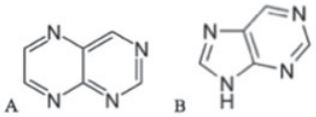

Figure 1. A-Pteridine: composed of a pyperazine nucleus fused to a pyrimidine cycle. Pteridines occur in pigments, in folic acid, in folates that are biochemical cofactors of group-transfer reactions. Pteridine is also involved in energy transfers. In the cell, pteridine is synthesized from GTP (guanosine triphosphate). B-Purine: composed of a pyrimidine cycle fused to an imidazole cycle.

must arbitrarily interrupt the reaction, because otherwise the reaction continues until compounds of high molecular weights and of defined structures are produced with total disappearance of the sugars of biological interest. Ribose is particularly unstable.

After having obtained the synthesis of the bases and sugars, prebiotic chemists naturally tried to bind these two compounds to produce nucleosides. Purine nucleosides are easily synthesized by heating and evaporating a mixture of adenine and ribose under vacuum at 100 °C [23–25]. Trying to use the same procedure to synthesize pyrimidine nucleosides occurs in very low yields (<0.1%). It is important to draw attention to the simplicity of these various reactions, a simplicity that is in their favour as initial candidates and as original sources of biochemical monomers.

Then, in the 1960s, Leslie Orgel spontaneously produced ribonucleic acids *in vitro* by template-directed synthesis starting from synthetic chemically-activated nucleotides. Jim Ferris did the same on clay surfaces of montmorillonite and it seemed at that time that we were close to the goal that was to obtain the first informational and replicative molecule in prebiotic conditions [17, 26]. However, the difficulties were pushed further away. Indeed how could one obtain RNA when the ribose, a part of the nucleotide, thus with a nucleic acid backbone, obtained through the formose reaction, is so unstable? How can this be liberated from the chemical synthetic-activation of the nucleotides essential for their polymerization?

Recently, convincing experiments on the way towards primordial synthesis of RNA polymers have been carried out by Rafaele Saladino starting from formamide, a common prebiotic compound [27]. John Sutherland succeeded in self-assembling activated pyrimidine ribonucleotides [28] and Steven Benner has focused on the stability of ribose by minerals such as ulexite, and coelemanite that are borate minerals [29]. In the same line of investigation, Thomas Georgelin stabilized ribose by silica [30] and Laura da Silva demonstrated the simple synthesis of RNA-like polymers by polymerization of standard nucleotides in the presence of ammonium chloride in hydrothermal conditions [31, 32].

These experiments might be considered as a what-if scenario of a pre-RNA world, namely a world in which an alternative genetic system (AGS), preceeded the advent of modern RNA molecules.

Another species of RNA-like polymers has been mimicked at the bench, for instance, peptide nucleic acids (PNA), that is peptides with an N(2-aminoethyl glycine) backbone bearing nitrogenous bases. Others have proposed synthetic genetic polymers such as pyranosyl RNA (p-RNA), Hexitol Nucleic Acid (HNA) or Altritol Nucleic Acids (ANA) or Threose Nucleic Acids (TNA), or synthetic polymers named XNA as mimosis models [33–35] (Fig. 2).

Nobody has claimed that such molecules were true remains of ancestral genetic molecules, but it is interesting to note that they are good templates in template-directed synthesis, meaning that a transition is possible between two different systems without loss of information [36], hence the suggestion that the first organisms might have had different informational systems, a kind of heterogenetic symbiotic organisms. Furthermore and in support to this idea, N(2-aminoethyl glycine) of PNA has been found in diverse taxa of cyanobacteria indicating that this molecule arose in early life [37].

Indeed, one of the most important yet unanswered questions is to understand how molecules as different as nucleic acids and proteins could have been formed spontaneously and simultaneously.

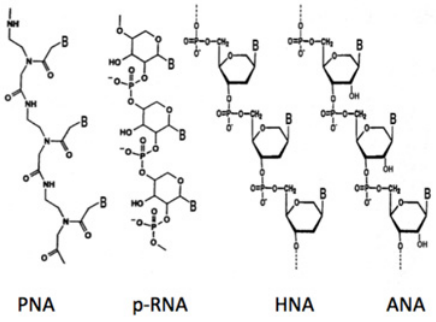

Figure 2. Examples of some alternative genetic systems (B = nitrogenous base).

One must recall that it was long thought that in the living cell, there exists a total separation between the roles of nucleic acids, DNA and RNA on the one hand that carry and transmit information, and on the other hand, proteins that as enzymes, structural proteins or signals, and that express this information in the cell. When in the 1980's Thomas Cech (1981) and Sidney Altman (1983) showed that certain classes of RNAs (ribozymes) found in introns perform catalytic reactions, this raised the question of the primitive catalysts used at the origins of life [38, 39]. Were the first autoreplicating molecules related to our nucleic acids also endowed with catalytic properties?

This hypothesis first put forward in 1967 by Carl Woese [40] was then developed in 1968 by Francis Crick [41] and Leslie Orgel [42]. They presented the model wherein RNAs would have been the first molecules of the living world capable of stocking genetic information, of auto-replication and of perfoming catalytic reactions!

3.1 Were the first nucleic acids also efficient catalysts?

By chemical modifications, nucleotides can acquire all the functional groups that amino acids possess in proteins with the sole exception of imidazole (Table 1 below).Yet imidazole is a constituant of histidine, an amino acid very frequently present at the active site of protein enzymes. The Table 1 presents the various functional groups frequently encountered on the one hand in the nucleotides of ribonucleic acids, and on the other hand in the amino acids of proteins. When the functional groups are absent from the initial monomers, the macromolecules can acquire them, either by post-transcriptional modifications, or by the addition of cofactors. This is possibly how biochemical evolution was able to proceed. Moreover, one can suppose that primitive nucleotides were not necessarily restricted to the standard nucleotides we know today.

3.2 From simple nucleoside catalysts to ribonucleotides

A simple, presumably prebiotic nucleoside, N6-ribosyl adenine behaves as a histidine analogue during the hydrolysis of para-nitrophenylacetate, thanks to the availability of the imidazole group, whereas in bases, standard purine nucleosides and nucleotides, the nitrogen (in 9 position) of the imidazole is "blocked" in an N-glycosidic bond (Fig. 3) [43].

As seen in the Fig. 3, when the ribose is bound in position 6 in N6-ribosyl adenine, rather than in position 9 as in standard adenosine, the imidazole ring is free and available for catalysis. Nucleotide analogues have also been tested during template-directed synthesis in the course of studies on primitive replication. It appears that 3-isoadenosine 5′-phosphate is easier to polymerize than the standard

Table 1. Functional groups common to RNAs, proteins and their cofactors.

Protein	RNA
Building blocks	
Hydrophilic group Hydrophobic group Sugars Amine Carboxylates Hydroxyl group Imidazole	Hydrophilic group Hydrophobic group Phosphates Sugars
Post-transcriptional modifications	
   Phosphates Sugars Cetone Selenium	Amine Carboxylates Sugars Hydroxyl group Cetone Selenium
Cofactors	
Nicotinamide Flavine Sulfonium ions	Nicotinamide R-SH Flavine Sulfonium ions

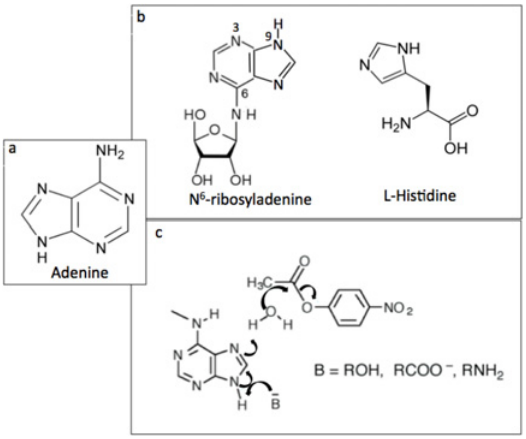

Figure 3. a) Adenine; b) comparison of modified adenosine and histidine; c) catalytic activity of adenine residue.

nucleotide when facing polyuridylic acid (polyU). Hoogsteen-type pairing involves positions 6 and 8 of the purine [44].

The original purine-purine pairings are also of interest in view of understanding a primitive RNA world. W´achtersh´auser (1988) proposes this type of original pairing because of the presence in numerous cell types of N3-ribosyl xanthine whose function remains unknown [45]. The nucleotide is synthesized in the cell from xanthine and phosphoribosyl thanks to uridine-pyrophosphorylase, an enzyme specific of pyrimidines. This reaction could represent the vestige of a purine-pyrimidine relationship, purine nucleotides being the precursors. Finally, this mode of pairing maintains the purine imidazole free, available for catalysis.

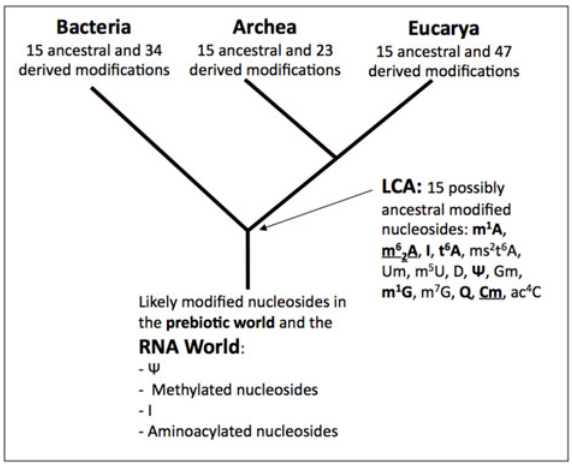

Figure 4. Phylogenetic analysis of ∼100 modified nucleosides. Nucleoside modifications might have occurred in the LCA and in the RNA world [47].

The biochemical synthesis of histidine is a critical path (its biosynthesis requires a great amount of energy: up to 41 ATPs) and yet its presence on the active site of numerous proteic enzymes is indispensable. Finally, histidine is difficult to obtain in prebiotic conditions since its cellular biosynthesis begins with a purine. This suggests that a primitive metabolic link exists between the two compounds. Starting from all these observations, we suggested [46] a biochemical scenario whereby catalytic groups present in past-day primitive nucleic acids became incorporated during the course of evolution into specific amino acids specialized in a specific type of catalysis. Thus, the purine N6 and N3 substituted-purines that possess replicative and catalytic properties could have been essential links between the world of nucleic acids and the world of proteins.

Furthermore, from the phylogenetic analysis of almost 100 modified nucleosides, it appears that nucleoside modifications might have occurred in the Last Common Ancestor (LCA) of living organisms and in the RNA world [47] (Fig. 4).

From their structure and chemistry they could have been present in great concentrations on the primitive earth. In addition nicotinamide that is the chemical moiety of the coenzyme nicotinamide adenine dinucleotide (NAD) has been synthesized in prebiotic conditions from ethylene and ammonia in the 1970's [20]. In the same line, present-day coenzymes, indispensable cofactors of many proteins would represent vestiges of catalysts of primitive metabolism and it is well accepted that ribonucleotides might be the remains of ancient RNAs [43, 48, 49].

4. RNA back and forth...

We know today that RNA performs multiple functions, including genetic, catalytic, structural and regulatory roles, supporting the fact that RNA came before the contemporary DNA, RNA and protein worlds.

Among the current cellular facts in favour of this possibility is that RNA is the indispensable primer for the replication of DNA, that thymidylic nucleotide (specific of DNA sequences) results fom methylation of uridylic nucleotide (specific of RNA) which come first in the modern cellular metabolism. We know numerous ribonucleotidic cofactors that are indispensable for protein enzyme activities, as well as the prominent role of ATP, the central ribonucleotide at work in all energetic actions, and also GTP, ADP and cAMP as metabolites for a range of living processes.

For over 30 years, discoveries have strengthened what is now regarded as the RNA world paradigm, and this is of great importance in the frame of the origins and early evolution of life to better understand the formation of the first RNA molecules.

DNA replication that always starts with the synthesis of ribonucleotide primers could correspond to modified transcription: RNA polymerisation would have been "displaced" during evolution by that of DNA. As a double-stranded molecule, DNA is indeed more stable than simple-stranded RNA. The absence of a hydroxyl –OH group in 2′ of the sugar, deoxyribose, is an additional stabilization element that must have contributed over the ages, to select DNA in its function of stocking genetic information. In the event that RNA would have appeared before DNA during biochemical evolution, DNA can be considered as a modified RNA. This take-over would simply be the logical next step of an evolutionary process during which other molecules would have preceded RNA and transmitted hereditary information.

Finally, the idea that RNA might have preceded DNA was strengthened by the existence of ribozymes which definitely confirmed the RNA world era [50].

Our knowledge of the *ribozyme* world (a contraction of the terms *ribo* and *enzyme*) as catalytic RNA, allows us to distinguish several types of catalytic activities [51].

Many RNA viruses, most of which are plant viruses possess ribozyme activity that participates in virus multiplication. One of the first to have been demonstrated is that of the satellite RNA of *Tobacco ringspot virus* (sTRSV). It is composed of 4 helices and 2 loops that come into close contact during catalysis occurring at the scissile phosphate between G and A in the first loop (Fig. 5). Ribozyme auto-cleavage activities were then demonstrated among plant viroids that are naked RNAs (lacking envelope and capsid) composed of circular RNAs, among virusoids that are circular satellite RNAs, among linear satellite RNAs, and in Hepatitis delta virus (HDV) that is associated with Hepatitis B virus. These 50 to 100 nucleotide-long ribozymes are presently the most studied models in view of determining the structural bases of RNA catalysis. It has been proposed [52] that viroid RNAs are vestiges of an ancient RNA world.

Three small ribozyme-like motifs are known with highly conserved secondary structures that are responsible for cleavage activity. These well-studied ribozymes are the hairpin motif, the hammerhead motif and the motif of HDV. Viroids, virusoids and certain linear satellites contain a common structure of about 30 nucleotides known as "hammerhead structure" whose T-shaped secondary structure contains various specific base-pairings that delineate the cleavage site. Other linear satellites possess a "hairpin" structure. The RNA of HDV adopts a pseudo-knot structure. The ribozyme of HDV that can perform auto-cleavage 100 times faster than a hammerhead ribozyme is therefore the most rapid natural ribozyme known. It is very stable, its optimal cleavage temperature is 65 °C and it is active up to 80 °C. It seems that divalent cations do not play a direct role in the catalysis and that cytosine 75, one of the bases of the catalytic site is directly involved in an acid-base type mechanism.

In certain linear RNA satellites of plant viruses, the hairpin ribozyme performs a reversible auto-cleavage reaction involved in the maturation of viral replication products. This ribozyme is composed of two helices linked to another one, each containing an internal loop forming two domains. When these two domains are side by side, they promote tertiary interactions constituting the active structure. As in the case of HDV, the bases of the ribozyme are the major catalytic determinants.

All these ribozymes undergo auto-cleavage at a unique specific site, but use different catalytic strategies. The sequence of these natural ribozymes can be modified to obtain RNAs performing inter-molecular cleavages or *trans* cleavages that consequently behave as true enzymes performing multiple turn-over reactions.

Two new "hairpin" ribozymes were discovered in our laboratory. They present only four to six mutations with respect to the wild-type sTRSV ribozyme. These ribozymes that utilise adenine as enzymatic cofactor are *coribozymes*, meaning that they depend on adenine to perform their reversible auto-cleavage reaction (Fig. 5) [53]. We studied in details and in "extreme" conditions the behaviour

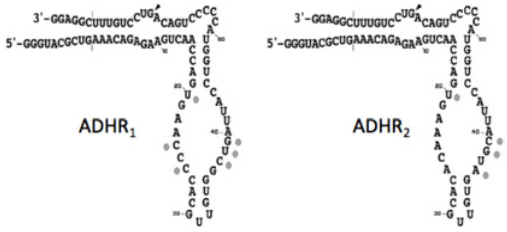

Figure 5. Adenine-dependent hairpin ribozymes (ADHR). Arrowheads: cleavage sites; grey dots: degenerated (mutated) sites; vertical bars: separation between the primer binding region and the random sequence.

of these ribozymes, demonstrating the extraordinary plasticity of these small RNAs and the diversity of their mechanisms in relation to various environmental conditions [54–56].

The year 2000 was the year of a major discovery: the 23S RNA contained in the large ribosomal subunit is sufficient to catalyze the formation of the peptide bond during protein synthesis. The ribosome is a ribozyme [57]. Peter Moore and Thomas Steitz who demonstrated that the ribosome is a ribozyme, described the cristallographic resolution of the large ribosomal subunit at the atomic level. The active site where the peptide bond is formed, is composed entirely of RNA, and is deprived of all protein elements. These latter elements positioned on the exterior of the cleavage site, only maintain the entire structure together.

To produce a protein, what is therefore required, is the information contained in the messenger RNA, a tRNA and a ribosomal RNA that will bind successively two amino acids!

Consequently, one now considers that ribosomal RNA derives from a proto-ribosome [58], a relic of the first genetically encoded systems of peptide synthesis. The selected peptides would first have helped to stabilize the RNA in its functions and make it reliable. This is how, according to White [48], the use by present-day proteins of ribonucleotidic cofactors can best be explained. During evolution, the protein moiety of the catalyst would have become structurally and functionally increasingly important and the ribozyme moiety would have progressively regressed, leaving only the ribonucleotidic cofactor.

In summary, *in vivo* RNA catalyzes trans-esterification reactions as well as the formation of peptide bonds during protein synthesis. *In vitro*, the performances of the RNA have been extended to various metabolic reactions, such as oligonucleotide linkages, polymerizations, transfer of aminoacyl groups, oxydative cleavages, etc.

The structural, and functional diversity of current RNAs is unexpected. rRNAs, tRNAs, mRNAs, RNase P, and the RNA of the signal recognition particle (SRP) are RNAs whose evolutionary history is the best understood [59]. It has even been shown that these molecules were present in the last common ancestor of all living organisms (LCA) [60]. For example, SRP is a ribonucleoprotein particle whose RNA moiety (SRP RNA or 7SL RNA) is highly conserved among eubacteria, archea and eukarya. Domains III and IV of 7SL RNA can base-pair forming a tRNA-like structure that interacts with the ribosome, causing a pause during translation. It was phylogenetically shown that, SRP and its RNA were present in the last common ancestor.

Other RNAs, the non-coding RNAs (ncRNAs), long or short RNAs such as nucleolar RNAs (snoRNAs), nuclear RNAs (snRNAs), vault RNAs which play a role in nucleocytoplasmic transport, micro RNAs (miRNAs) and siRNAs also seem to have originated very early, but their more complex evolutionary history still remains to be discovered. Another oddity of the modern RNA world is TmRNA (Transfer-messenger RNA) [61], a stable cytoplasmic RNA found in eubacteria, but unknown among archea and eukarya. TmRNAs possess a "t^{Ala}RNA-like" structure and an internal coding region that codes for a short signal peptide. It performs a new type of translation, *trans-translation*, during which a peptide is produced from two distinct mRNAs. The TmRNA acts both as tRNA and as mRNA to "unblock" ribosomes stalled on an mRNA lacking a stop codon. The aminoacylated TmRNA then

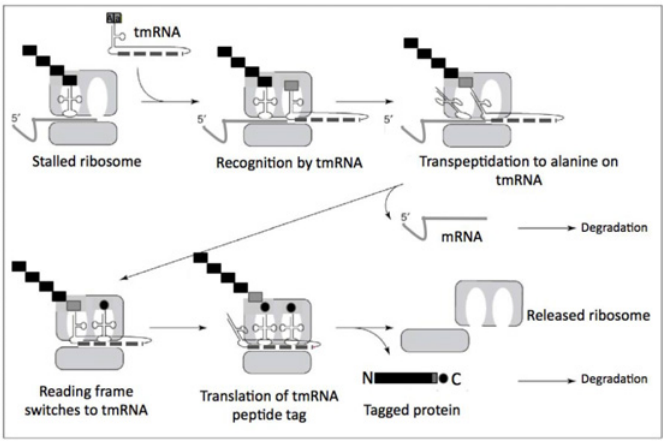

Figure 6. Tm RNA (from [62]).

comes into play and adds alanine to the growing peptide chain. The ribosome releases the truncated mRNA and begins to translate the reading frame of the TmRNA. A chimeric protein is thus produced whose C-terminal signal will be recognized by proteases that will degrade this small peptide signal. This control mechanism of translation is important for the bacterial cell since the "blocked" ribosome now becomes available for other translation cycles. Does the TmRNA that has a double function (just as $tRNA^{Ala}$ it can be charged by the corresponding alanyl-tRNA synthetase, and as mRNA its reading frame can be translated by the ribosome) correspond to a bacterial adaptation, or was it lost during the course of evolution by early organisms and/or by the archea and eukarya? (Fig. 6).

Another less-known competence of RNA results from its fragmentation as cellular stress responses. RNA fragments derived from well-characterized parent RNAs (mRNAs, snoRNAs etc.) that perform functions distinct from those of their parents, for instance binding other transcripts, acting as trans-silencers, or changing a repressor into an activator [63]. Hence the idea that thanks to its plasticity, to its capacity to exert new functions after breaks and reorganizations, RNA is at the origin of evolutionary innovations.

It is not possible to describe here the impressive list of the numerous RNAs discovered in recent years and that participate in present-day metabolism, as catalysts, regulators, circulating RNAs, antiviral defence agents, or actors of cell development....

Epi-transcriptomic in response to environmental modifications intervenes in dynamics and morphogenetic events on the path towards the construction of organisms from simple ones to the more complex. Our research that takes into account the molecular, environmental, historical and biological aspects, resides within the realm of acheobiology. Thus one can propose that what we are discovering today is, but the pieces of an immense biological "iceberg", that shuttles back and forth between the ancient and the modern RNA world.

The assistance of Anne-Lise Haenni and of Laura da Silva in preparing the manuscript is gratefully acknowledged.

References

[1] K. Popper, *Unended Quest: An Intellectual Autobiography* (Routledge Classics, 1982)
[2] A.I. Oparin, *Proiskhozhdenie Zhizny [The origin of life]* (Izd. Moskvskiy Rabochiy, 1924)

[3] R.J. Horodyski, L.P. Knauth, Science **263**, 494 (1994)
[4] S. Gruber, J.W. Veening, J. Bach, M. Blettinger, M. Bramkamp, J. Errington, Current biology, **24**, 3, 293–298 (2014)
[5] M-C Maurel, B. Brun, Mésogée, **57**, 3–10 (1999)
[6] B. Brun, M-C Maurel, *L'inné et l'acquis nouvelles approches épistémologiques*, 7–46 (publications de l'Université de Provence, 2005)
[7] Lucrece, Of Nature, translated by A. Ernout (Société d'Edition "les Belles Lettres", Paris, 1924)
[8] F. Wöhler, Annales de Chimie et de Physique, **37**, 330–334 (1828)
[9] A.M. Butlerov, Zeitschrift für Chemie, **4**, 549–560 (1861)
[10] W. Lôb, Berichte der deutsche chmischen Gesellschaft, **46**, 684–697 (1911)
[11] D. Davidson and O. Baudisch, J. American Chemical Society, **48**, 2379 (1926)
[12] The correspondence of Charles Darwin, Letter to J.D. Hooker, 19 (1871)
[13] S.L. Miller, Science **117**, 3046, 528–529 (1953)
[14] J. Oro, Nature **191**, 1193–1194 (1961)
[15] J. Oro, T. Mills, A. Lazcano, Orig. Life Evol. Biosph. **21**, 267–277 (1992)
[16] J. Ferris, L.E. Orgel, J.A.C.S, **88**, 5, 1074 (1966)
[17] J. Ferris, J. Kuder, A. Catalano, Science, **166**, 3906, 765–766 (1969)
[18] M. Berthelot, *La synthèse chimique* (Felix Alcan éditeur, Paris, 1887)
[19] E. Desguin, A. Dauvillier, *La genèse de la vie, phase de l'évolution géochimique* (Hermann, Paris, 1942)
[20] N. Friedman, S.L. Miller, R.A. Sanchez, Science, **171**, 3975, 1026–1027 (1971)
[21] M.S. Kritsky, T.A. Telegina, Y.L. Vechtomova, M.P. Kolesnikov, T.A. Lyudnikova, and O.A. Golub, Biochemistry (Moscow), **75**, 10, 1200–1216 (2010)
[22] R. Breslow, Tetrahedron Letters, **21**, 22–26 (1959)
[23] W.D. Fuller, R.A. Sanchez, L.E. Orgel, J. Mol. Biol. **67**, 25–33 (1972)
[24] W.D. Fuller, R.A. Sanchez, L.E. Orgel, J. Mol. Evol. **1**, 249–257 (1972)
[25] M-C. Maurel, O. Convert, Origins of Life Evol. Biosphere, **20**, 43–48 (1990)
[26] T. Wu, L.E. Orgel, J. Am. Chem. Soc. **114**, 317–322 (1992)
[27] R. Saladino, C. Crestini, G. Costanzo, R. Negri, E. Di Mauro, Bioorganic and Medicinal Chemistry. **9**, 5, 1249–1253 (2001)
[28] M.W. Powner, B. Gerland and J.D. Sutherland, Nature **459**, 239–242 (2009)
[29] A. Ricardo, M.A. Carrigan, A.N. Olcott, S.A. Benner, Science **303**, 196 (2004)
[30] T. Georgelin, M. Jaber, F. Fournier, G. Laurent, F. Costa-Torro, M-C. Maurel, J-F. Lambert, Carbohydrate Research, **402**, 241–244 (2015)
[31] L. Da Silva, M-C. Maurel, D. Deamer, J. Mol. Evol (In press)
[32] B.J. Cafferty, N.V. Hud, Current Opinion in Chemical Biology, **22**, 146–157 (2014)
[33] I.A. Kozlov, M. Zielinski, B. Allart, L. Kerremans, A. Van Aerschot, R. Busson, P. Herdewijn, L.E. Orgel, Chemistry. **6**, 151–155 (2000)
[34] I.A. Kozlov, B. De Bouvere, A. Van Aerschot, P. Herdewijn, L.E. Orgel, J. Am. Chem. Soc. **121**, 5856–5859 (1999)
[35] J. Attwater, P. Holliger, Nature Methods **11**, 495–498 (2014)
[36] G.F. Joyce, A.W. Schwartz, S.L. Miller, L.E. Orgel, Proc. Natl. Acad. Sci. **84**, 4398–4402 (1987)
[37] S.A. Banack, J.S. Metcalf, L. Jiang, D. Craighead, L.L. Ilag, P.A. Cox, PLOS ONE, **7**, 11, e49043 (2012)
[38] T.R. Cech, A.J. Zaug, P.J. Grabowski, Cell **27**, 487 (1981)
[39] C. Guerrier-Tanaka, K. Gardiner, T. Marsh, N. Pace, S. Altman, Cell **35**, 849–857 (1983)
[40] C.R. Woese, Proc. Natl. Acad. Sci. **54**, 1546–1552 (1965)
[41] F.H. Crick, J. Mol.Biol. **38**, 367–379 (1968)
[42] L.E. Orgel, J. Mol. Biol. **38**, 381–393 (1968)

[43] M-C. Maurel, J. Ninio, Biochimie **69**, 551–553 (1987)
[44] A.R. Hill, S. Kumar, N.J. Leonard, L.E. Orgel, J. Mol. Evol, **27**, 2, 91–94 (1988)
[45] G. Wachtershauser, Proc. Natl. Acad. Sci. **85**, 1134–1135 (1988)
[46] M-C. Maurel, J. Evol. Biol 5, 173–188 (1992)
[47] N. Cermakian, R. Cedergren, *Modification and editing of RNA*, 535–541 (ASM Press, Washington DC, 1998)
[48] H.B. White, J. Mol. Evol. **7**, 101–104 (1976)
[49] M-C. Maurel, J-L. Décout, J. Mol. Evol. **35**, 190–195 (1992)
[50] W. Gilbert, Nature **319**, 618 (1986)
[51] R. Hernández-Morales, A. Becerra and A. Lazcano (In press)
[52] T.O. Diener, Proc. Natl. Acad. Sci. **86**, 9370–9374 (1989)
[53] M. Meli, J. Vergne, M-C Maurel, J. Biol. Chem. **278**, 11, 9835–9842 (2003)
[54] S. Tobé, T. Heams, J. Vergne, G. Hervé, M-C Maurel, Nucleic Acids research, **33**, 8, 2557–2564 (2005)
[55] M. Ztouti, H. Kaddour, F. Miralles, C. Simian, J. Vergne, G. Hervé, M-C Maurel, FEBS J. **276**, 9, 2574–2588 (2009)
[56] Y.L. Li, M-C. Maurel, C. Ebel, J. Vergne, V. Pipich, G. Zaccai, Eur Biophys J. **37**, 2, 173–182 (2008)
[57] T.R. Cech, Science **289**, 878–879 (2000)
[58] C. Davidovich, M. Belousoff, A. Bashan, A. Yonath, Research in Microbiology **160**, 487e492 (2009)
[59] Meli, M., Albert-Fournier, B., and Maurel, M-C. International Microbiology, **4**, 5–11 (2001)
[60] J.F. Atkins, R.F. Gesteland, T.R. Cech, *RNA worlds: From Life's Origins to Diversity in Gene Regulation* (Cold Spring Harbor, NY:Cold Spring Harbor lab. Press, 2011)
[61] B. Felden, H. Himeno, A. Muto, J.F. Atkins, R.F. Gesteland, Biochimie **78**, 979–983 (1996)
[62] K.C. Keiler, Current Opinion in Microbiology, **10**, 169–175 (2007)
[63] C. Tuck, D. Tollervey, Cell. **27**, 10, 422–431 (2011)